



Available online at <http://scik.org>

Commun. Math. Biol. Neurosci. 2022, 2022:1

<https://doi.org/10.28919/cmbn/6838>

ISSN: 2052-2541

## IMPROVING LUNG DISEASE DETECTION BY JOINT LEARNING WITH COVID-19 RADIOGRAPHY DATABASE

HERY HARJONO MULJO<sup>1,\*</sup>, BENS PARDAMEAN<sup>2,3</sup>, KARTIKA PURWANDARI<sup>3,4</sup>,

TJENG WAWAN CENGGORO<sup>3,4</sup>

<sup>1</sup>Information Systems Department, School of Information Systems, Bina Nusantara University, Jakarta 11480,  
Indonesia

<sup>2</sup>Computer Science Department, BINUS Graduate Program, Bina Nusantara University, Jakarta 11480, Indonesia

<sup>3</sup>Bionformatics and Data Science Research Center, Bina Nusantara University, Jakarta 11480, Indonesia

<sup>4</sup>Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia

Copyright © 2022 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract:** Diagnostic chest radiography is one of the most common imaging tests performed in medical practice. A radiology workflow goal is to detect, diagnose, and manage diseases using chest radiography in an automated, timely, and accurate manner. Radiography data have proved very effective for assessing COVID-19 patients, particularly for treating overcrowded emergency departments and hospitals. The use of Deep Learning (DL) methods in Artificial Intelligence (AI) has become dominant in detecting diseases via chest X-rays. This study utilized the COVID-19 Radiographic Database and the National Institutes of Health (NIH) Chest-Xray to study pre-training fine-tuning of the DL model on chest radiographic images. We investigate the robust network architecture in detail: DenseNet-121, in this dataset dual technique to improve insight into the different methods and their application to chest X-ray

---

\*Corresponding author

E-mail address: [heryhm@binus.edu](mailto:heryhm@binus.edu)

Received September 29, 2021

classification. Consequently, this dual dataset technique is able to provide better detection results for each cluster of lung diseases. AUC results obtained using DenseNet-121 reached an average of 82.16 percent, with the highest AUC reaching 99.99% in the cluster containing Viral Pneumonia lung disease.

**Keywords:** chest radiography; lung disease; deep learning; image augmentation; transfer learning.

**2010 AMS Subject Classification:** 93A30, 65D18.

## 1. INTRODUCTION

Corona Virus Disease 19 (COVID-19) [1]–[3], a disease that has become a serious health concern for the entire world, can cause respiratory problems, heart infections, and even death [4]. In response to intensified travel between countries, the World Health Organization (WHO) declared the disease a global pandemic on March 11, 2020 [5], [6]. COVID-19 causes lungs to deteriorate and mutate before treatment is administered to patients based on a diagnosis. Therefore, lung examinations by radiologists are important for people who are experiencing COVID-19 complications [7], [8]. Only 1,578 of the approximately 41,000 specialist doctors in Indonesia are radiologists, according to the chairman of the MKKI IDI, David S. Perdanakusuma in 3,000 hospitals and 10,000 community health centers [9]. Nevertheless, there are still discrepancies in the distribution of specialists with the provinces of Papua, Maluku, Nusa Tenggara, Sulawesi, and DIY, North Sulawesi, reporting the highest disparity levels. Therefore, there are not as many specialists as there are hospitals in Indonesia [10].

A diagnostic tool based on advanced technology is absolutely necessary for today's era [11]. Physicians who specialize in radiology are known as radiologists and use imaging techniques to detect, diagnose, and treat disease by using imaging procedures, like X-rays, CT scans, magnetic resonance imaging (MRI), nuclear medicine, and ultrasound [12]. Boosting the number of radiologists is crucial to solving the problem of the increased technology of medical devices without also increasing the number of radiologists. This is especially true since radiology is often crucial for diagnosis [13], [14].

Worldwide, approximately one billion radiation studies are conducted each year [15]. Additionally, physicians are often in agreement that medical imaging should be led by radiology

[16]. In this way, the right diagnosis is important for the success of therapeutic or treatment action, since misdiagnosis can lead to ineffective treatments or even death if the diseases are not correctly diagnosed.

A solution to the uneven distribution of specialists in Indonesia can be using artificial intelligence (AI) technology in hospitals [17]. Recent decades have seen technological advances in healthcare [15]. It has been demonstrated that AI is capable of improving healthcare delivery and management, clinical decision support, and personalized medicine [16]. A possible future application of AI in medicine is working alongside doctors [18].

Additionally, AI can help reduce and eliminate factors that cause a human error, such as fatigue [19], and increases the accuracy of doctors' diagnoses. To assist radiologists, a lung disease detection information system with AI technology has been created. A doctor can use this information system to support their diagnosis by automatically detecting lung disease. In addition to using the results to determine which patients are worth a more thorough examination, the results will also be employed as an initial screening tool for doctors. A deep learning [20] model was developed for increasing the accuracy and consistency of early detection [18].

Infections caused by COVID-19 can be quickly assessed using medical images and artificial intelligence (AI) [21]. Hence, with limited data, an AI technique is urgently needed to classify COVID-19 images in a short time period. Its power for image classification has been demonstrated with superhuman accuracy and depth by deep learning. The availability of large datasets with quality ground-truth annotations is one major challenge in the medical domain. This problem can often be overcome by the transfer learning approach [13], [22], [23].

A chest radiograph classification task with two datasets is assessed as part of this paper. Image-based classification is the only measure we use to compare algorithms and radiologists. DenseNet121 [24], a deep Convolutional Neural Network (CNN) [25] architecture, was trained and validated on training and validation sets, and then evaluated on the test set based on the attending radiologists' labels. The performance of the model was evaluated using receiver operating characteristic (ROC), AUC, and precision-recall (PR) curves, and confusion matrix analysis. In this study, we present a relatively simple but robust performance finding of

DenseNet121 that could enhance the accuracy of classifying COVID-19 Radiographs and the National Institutes of Health (NIH) Chest X-Ray Database using minimal training.

## 2. RELATED WORK

The medical field has shown impressive performance with Deep Learning (DL) models [14], [26], [27]. Therefore, the main purpose of DL is to detect lung disease infections in multiple ways. Several researchers have employed CNN to speed up the analysis of lung disease infected images [18]. An artificial intelligence (AI)-based model was developed by using chest radiographs, followed by testing it on an independent cohort, and achieving 0.973 screening accuracy for pneumoconiosis and 0.927 staging accuracy for pneumoconiosis, respectively [28]. Systemic sclerosis patients with interstitial lung disease can be diagnosed using deep learning methods similar to those used by radiologists [29]. Deep learning-based algorithms can also be developed to detect chest radiographs that indicate major thoracic diseases and exhibit high and consistent performance [30]. A new hybrid deep learning framework called VDSNet was successfully used to detect lung diseases from X-ray images with a validation accuracy score of 73% without having to extend training time [31].

Initially, COVID-19 labeled datasets were suitable for practical implementation by downloading from open access repository [32], [33]. Chest X-Ray 14 is a dataset made publicly available that includes fifteen classes: Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis, Pleural Thickening, Hernia, and No Finding images. Machine learning and deep learning are used to analyze data and to create models for diagnosing patients [33]. Using the well-known pre-trained model for CNN, we can make the best prediction possible about patients based on their X-rays.

Refer. [26] describes comprehensive studies on the detection of consolidation using DenseNet121 and VGG 16 [34]. Clinically significant pulmonary masses/nodules can be detected on chest X-ray images by a deep learning-based Computer Aided Design (CAD) system. Furthermore, an approach based on deep learning is also described in Ref. [28] where several methods have been applied, such as DenseNet121, AlexNet [35], Inception V3 [36], etc., to

diagnose pneumonia [37]. Nevertheless, implementing the methods of these methods is extremely complex in terms of parameter tuning. The accuracy rate of DensNet-121 for lung lesions was 98.88%. With the traditional DenseNet-121 method for lung lesions, which uses ReLU and SGD, this accuracy rate is higher—81.06%, 81.14%, 80.74%, and 80.86% respectively. A traditional DenseNet-121 method produced false predictions for 8 images showing an image of a normal condition and 8 images showing an image of a lung disease. However, true predictions included 224 images of a normal condition and 340 images of lung disease [38].

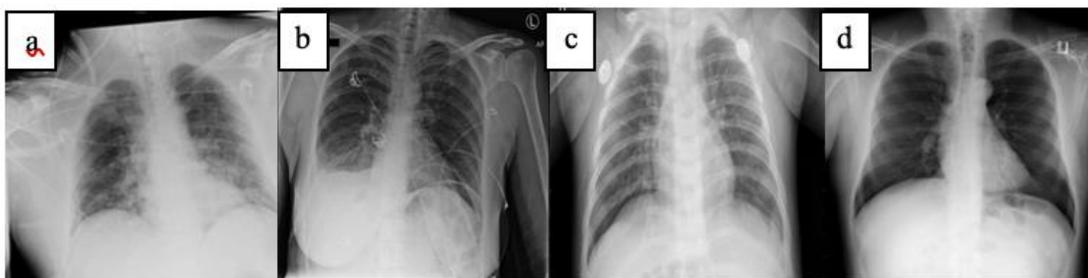
The data is not preprocessed or augmented because of the large training loss. It is based on an image augmentation model. Due to proper augmentation and weighted losses, the model is more accurate overall than our former best model [39].

### 3. MATERIAL AND METHODS

#### 3.1. Dataset

Our study used a combined dataset from two public databases (Kaggle) to classify images. Firstly, we provide another chest X-ray dataset named COVID-19 Radiography Database [40]. Through the Kaggle API, we were able to extract the data directly from Kaggle. In this dataset, 3616 COVID-19 positive cases are represented along with 10,192 normal and 6012 lung opacities (non-COVID lung infections), as well as 1345 cases of viral pneumonia. Figure 1 shows the example of chest X-ray images of COVID-19, lung opacity, viral pneumonia, and normal subjects provided by the COVID-19 Radiography Database.

FIGURE 1. Example of Chest X-Ray Images of (a) COVID-19, (b) Lung Opacity, (c) Viral Pneumonia, and (d) Normal.



In the second dataset, National Institutes of Health (NIH) Chest X-rays [41] have been analyzed according to 14 different lung conditions. A total of 112,000 chest X-ray images taken on more than 30,000 unique patients are available through NIH Chest X-rays. An example of a Chest X-ray obtained from the NIH Chest X-rays Dataset is shown in Figure 2.

FIGURE 2. Variation Chest X-Rays of Lung Disease from NIH Chest X-Rays Dataset



Our DenseNet121 model was trained on three categorizations of the dataset. First, the images were categorized using a combination of images taken from COVID-19 Radiography and the NIH Chest X-rays dataset (133,280 observations). Further, the second category consists of the original COVID-19 Radiography dataset (21,165 observations), while the third category pertains to the original NIH Chest X-rays dataset (112,115 observations).

### 3.2. Data Preparation

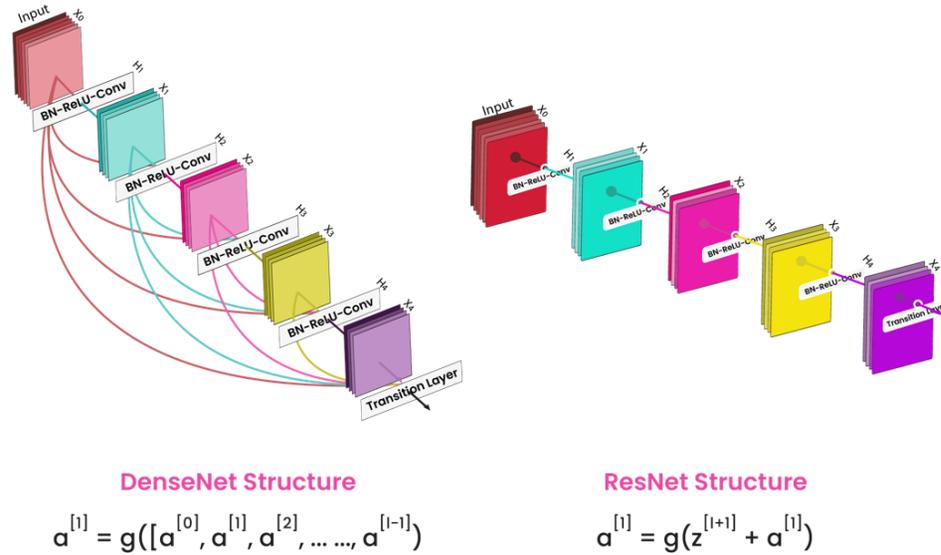
By altering the original training set image, augmentation contributes to increasing the diversity of the training data set. Ultimately, this allows the model to more accurately predict new images and expands the size of the training data set. This x-ray dataset is trained with augmentation data, including random horizontal flip, scaling of 224 pixels, center crop of 224, normalization with mean = [0.485, 0.456, 0.406] and standard deviation = [0.229, 0.224, 0.255]. In the meantime, the data validation and testing are done with the same augmentation data as the training data, but do not follow the random horizontal flip augmentation process.

Each process loads image data according to the fold specified in the dataset. During training, testing, and validation, every category of data is loaded by the data loader. A data loader tool makes it possible to load data regularly from the front.

### 3.3. Deep Learning

In this study, pre-training for DenseNet121 was selected. DenseNets are beneficial, as they reduce the missing gradient problem, strengthen feature propagation, stimulate feature reuse, and reduce the number of parameters. DenseNet121 transfer learning architecture and parameters are schematically depicted in Figure 3.

FIGURE 3. DenseNet Structure Concept



In a composite function operation, an output from one layer becomes an input from another. There are four components to this composite operation: a convolution layer, a pooling layer, a batch normalization layer, and a non-linear activation layer. Direct connections of this type are defined as having  $L(L+1)/2$  connections. In architecture,  $L$  is the number of layers. Several versions of DenseNet exist, such as DenseNet-121, DenseNet-160, or DenseNet-201. Numbers indicate the number of layers in a neural network. In order to compute 121, follow these steps:

$$\text{DenseNet-121: } 5 + (6+12+24+16)*2 = 121$$

5 – Convolution and Pooling Layer

3 – Transition Layers (6,12,24)

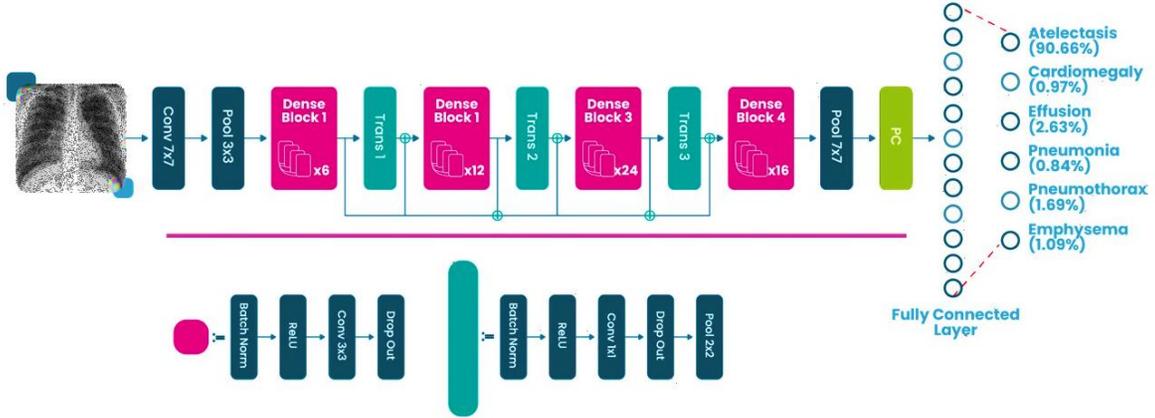
1 – Classification Layer (16)

2 – DenseBlock (1x1 and 3x3 conv)

In any case, the addition or concatenation of layers requires the same dimensions in the feature map. DenseNet consists of DenseBlocks, each with a number of filters, but dimensions within the

block are the same. Downsampling is applied in the Transition Layer to achieve batch normalization; it's an essential step in CNN. Figure 4 illustrates the DenseBlock inside and the transition within it.

FIGURE 4. DenseNet with DenseBlock and Transition Layer for Chest X-Ray Classification



By changing the filter count, the channel size increases between the DenseBlocks. It is important to generalize the first layer by using the growth rate ( $k$ ) as shown in equation (1). Essentially, it decides how much info to add to each layer.

$$k^{[l]} = (k^{[0]} + k(l - 1)) \quad (1)$$

X-ray images of the chest were converted to RGB images and finally resized so they would fit into each pre-trained CNN. For the training option, stochastic gradient descent with momentum optimizer is used, where momentum value = 0.9; weight decay = 0.0001; minimum batch size = 50; a maximum number of epochs = 100; early learning level = 0.01; constant learning rates are used throughout; randomizing the training and validation data is performed before each training period.

### 3.4. Evaluation Metrics

#### 1. Confusion Matrix

Confusion matrices and contingency tables are two types of structures used to explain classification decisions. There are four categories of confusion matrix: True positives (TP), which are example of positives classified correctly. Positive examples mistakenly labelled as negatives are known as false positives (FPs). A true negative (TN) is a negative that was

correctly classified. Finally, false negatives (FN) refer to positive examples that have been improperly labeled as negative. ROC space or PR space can be constructed by using the confusion matrix [42].

## **2. Area Under Curve (AUC)**

Generally, an AUC > 50% indicates a performance better than chance, while an AUC closer to 100% indicates a model that performs as well as possible. AUC values are generally used as a way to assess the accuracy of diagnostic tests. AUC values within the range of 0–1 can be considered as accurate [43].

AUC values that approach 1 indicate that the diagnostic test is improving. Based on the accuracy values interpreted in the table, there are five accuracy categories: values 50% - 60% (very weak), 60% - 70% (weak), 70% - 80% (moderate), 80% - 90% (strong), 90% - 100% (very strong). Accordingly, AUC values are: >50% - 60% (very weak), >60% - 70% (weak), >70% - 80% (moderate), >80% - 90% (good), >90% - 100% (very good). A positive predictive value is a measure of the likelihood that a person will develop disease if they test positive. People who test negative for disease have a negative predictive value, which indicates there is no disease present. Clinical interpretation of a test result requires the presence of positive predictive value. Values based on sensitivity and specificity are influenced by other factors as well, such as prevalence (prior probability), which varies with different situations. Additionally, likelihood ratios can be used as a measure of examination accuracy [44].

## **3. Receiver Operating Characteristic (ROC) Curve**

ROC analysis is used to describe, organize, and classify several categories in a statistical model based on their results. Through diagnostic testing, ROC curves are used in the medical field to analyze decision-making. The ROC graph displays the relationship between True Positive Rate (TPR), also known as Sensitivity (Y-axis), and False Positive Rate (FPR), also known as Specificity (X-axis). An unbiased probability value (0,0) on the ROC graph means that it never shows a positive result, meaning that the classification never produces false positives or true positives [45].

## **4. Precision-Recall (PR) Curve**

Precision-Recall Curve can be used as a method of drawing performance curves for classes

with positive attributes (which are typically found on minority scales). Using the confusion matrix, a formula is developed to produce the curve. This value is used to calculate precision and recall. When dealing with unbalanced classes, it is critical to focus on positive class performance. In PR, reaching the top right corner (1,1) is the goal. Accordingly, the top right corner indicates that 100% of the positives (Remember = 1) equal zero Positive False, and the true positives (Precision = 1) equal zero Positive True [46].

## **4. EXPERIMENTAL SETUP**

### **4.1. Dataset Splitting**

Pareto ratio-based testing was applied to 20% of the whole dataset. Another split of the remaining datasets has been done using Pareto principle for training and validation (80% is training, 20% is validation).

### **4.2. Pre-processing**

A small dataset usually overfits DL models. To achieve good generalization and effective training, there must be enough data. By incorporating multiple variations in the base dataset, data augmentation is an effective method of improving the generalization of the learning model. Different types of transformations are applied to the training samples in this study, including horizontal and vertical reflections, rotations, and shears. 224x224x3 images have been resized before being assigned to CNN for training.

### **4.3. Model Implementation Detail**

An end-to-end method was used to train deep CNN models. For reducing the cross-entropy loss, a stochastic gradient descent (SGD) algorithm was employed. Identifying class probabilities was achieved using Softmax. Piecewise learning rate scheduler was used for this training with an initial value of 0.0001 and a momentum of 0.95. On NVIDIA GPU Tesla P100, the training time for 1 epoch was 1 hours. For each of the three datasets, the implemented models were trained and evaluated on an unseen test set.

### **4.4. Working Environment**

Pytorch was used to build CNN models, and the DL library was used to simulate the models. The experiments were performed on a NVIDIA Tesla P100 computer which has 149 GB RAM, along with a CUDA enabled GPU.

## 5. RESULT

This lung classification was based on multi-class results. DenseNet121 predicts the results for each class by a probability. Based on this X-ray dataset, it can be said that the best classification result would be that with the highest probability. Table 1 displays that the "effusion" class received the highest results compared with the rest.

<b>Lung Disease</b>	<b>Predicted Probability</b>
Cardiomegaly	0.2747
Emphysema	0.0132
Effusion	<b>0.4216</b>
Hernia	0.2046
Infiltration	0.0336
Mass	0.0156
Nodule	0.0141
Atelectasis	0.0265
Pneumothorax	0.1044
Pleural_Thickening	0.0268
Pneumonia	0.0211
Fibrosis	0.0038
Edema	0.0195
Consolidation	0.0002
Viral Pneumonia	0.0009
Normal	0.0005
Lung_Opacity	0.0024
COVID	0.0042

TABLE 1. Probability prediction of combined X-ray dataset.

Using a combined dataset, Figure 5 shows the training processes of transfer learning. As illustrated in Figure 6, the training process of transfer learning applies to the CXR Dataset, while Figure 7 shows the transfer learning process using the COVID-19 Dataset. Learning curves calculated from a training dataset provide an indication of how well the model is learning, as well as those calculated from a hold-out validation dataset providing an indication of how well the model generalizes.

Typically, two learning curves are created during the training of a machine learning model on both the training and validation data. This shows that overfitting occurred when the DenseNet121 model was used for learning the data.

FIGURE 5. Learning Curve of CXR Dataset and COVID-19 Dataset

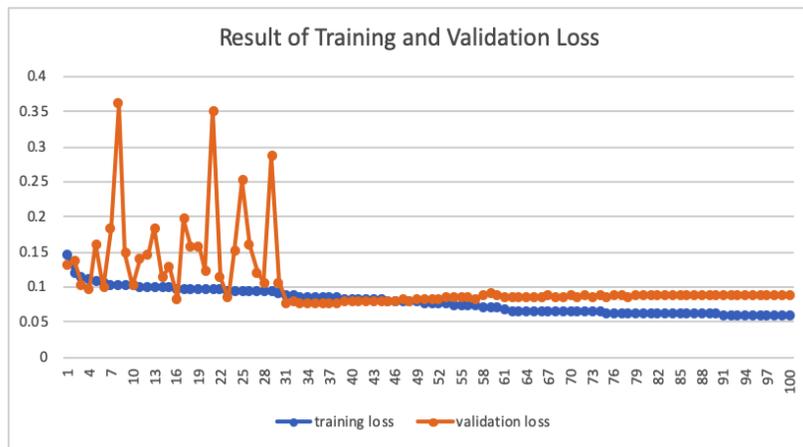
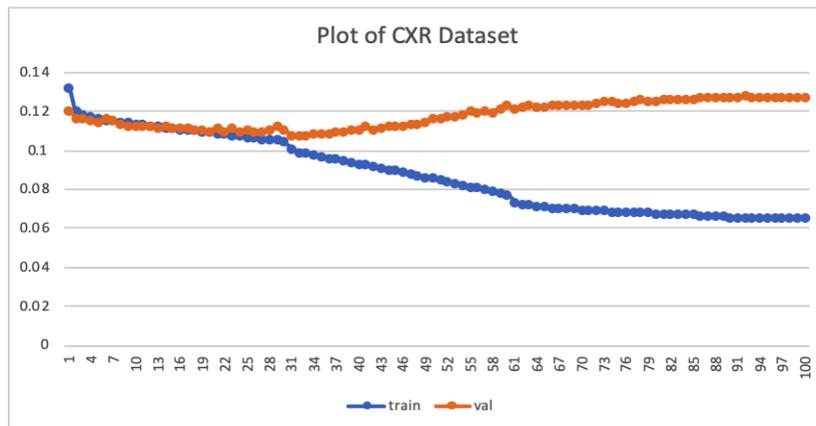


FIGURE 6. Learning Curve of CXR Dataset



## IMPROVING LUNG DISEASE DETECTION BY JOINT LEARNING

FIGURE 7. Learning Curve of COVID-19 Dataset

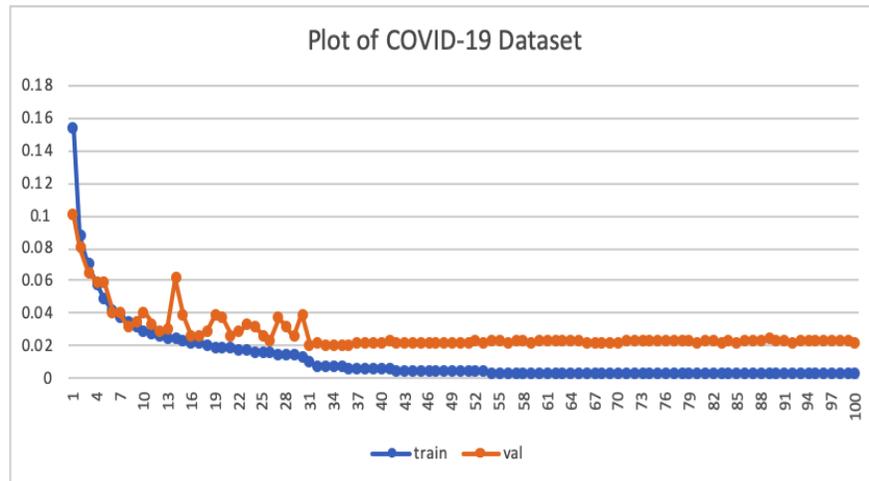


Table 2 shows the results of AUC in each dataset categorization. Results in most classes have significantly increased with the use of original NIH Chest X-rays dataset compared with a combination of both, with the exception of Cardiomegaly. Nevertheless, there have been increased examinations using a combination of COVID-19 and the radiography dataset in all subjects. Additional datasets with varied classes have been used to prove the DenseNet121 model's success.

Lung Disease	AUC of NIH Chest X-rays	AUC of COVID-19 Radiography Database	AUC of NIH Chest X-rays + COVID-19 Radiography Database
Atelectasis	0.733637	-	<b>0.754644</b>
COVID	-	0.999272	<b>0.999897</b>
Cardiomegaly	<b>0.837207</b>	-	0.829576
Consolidation	0.707358	-	<b>0.732540</b>
Edema	0.823728	-	<b>0.835363</b>
Effusion	0.796364	-	<b>0.812568</b>
Emphysema	0.773421	-	<b>0.817715</b>
Fibrosis	0.752721	-	<b>0.772040</b>
Hernia	0.742819	-	<b>0.791765</b>

Infiltration	0.671275	-	<b>0.712099</b>
Lung Opacity	-	0.983737	<b>0.999177</b>
Mass	0.718600	-	<b>0.757600</b>
Module	0.673160	-	<b>0.702578</b>
Normal	-	0.982174	<b>0.998693</b>
Pleural Thickening	0.705496	-	<b>0.739807</b>
Pneumonia	0.687467	-	<b>0.709697</b>
Pneumothorax	0.800954	-	<b>0.824159</b>
Viral Pneumonia	-	0.998409	<b>0.999947</b>

TABLE 2. Comparison of AUC of Three Dataset Categorizations Using 100 Epochs.

In both "Emphysema" and "Viral Pneumonia", the drain was chosen as the main feature. We have visualized the heatmap of the factors most responsible for the final prediction. As seen in Figure 8(a) and 8(b), the highest activations are found around the drain in the sample test set from "Emphysema" and "Viral Pneumonia", respectively. Hence, DenseNet121 has been trained to detect both an acute Emphysema and a Viral Pneumonia as well as chest drains. Typically, our proposed model can be found at the location of the lung abnormality that warrants a diagnostic assessment. It is located in the part of the lung that changes from white to black when a patient has emphysema. Those suffering from Emphysema typically suffer from swelling of small air sacs within the lungs (alveoli). Emphysema conditions cause more than just swelling to the alveoli, as their elasticity can be lost. Additionally, Emphysema can also be classified as an obstructive chronic lung disease [47]. A viral pneumonia disorder occurs when the lungs become infected with certain microorganisms, and the alveoli that should be filled with air become filled with fluid or pus. Fluid buildup can cause white patches on the lungs in several locations as a result of lung changes [48].

## IMPROVING LUNG DISEASE DETECTION BY JOINT LEARNING

FIGURE 8. Result of Grad-CAM from Emphysema (a) and Viral Pneumonia Class (b)

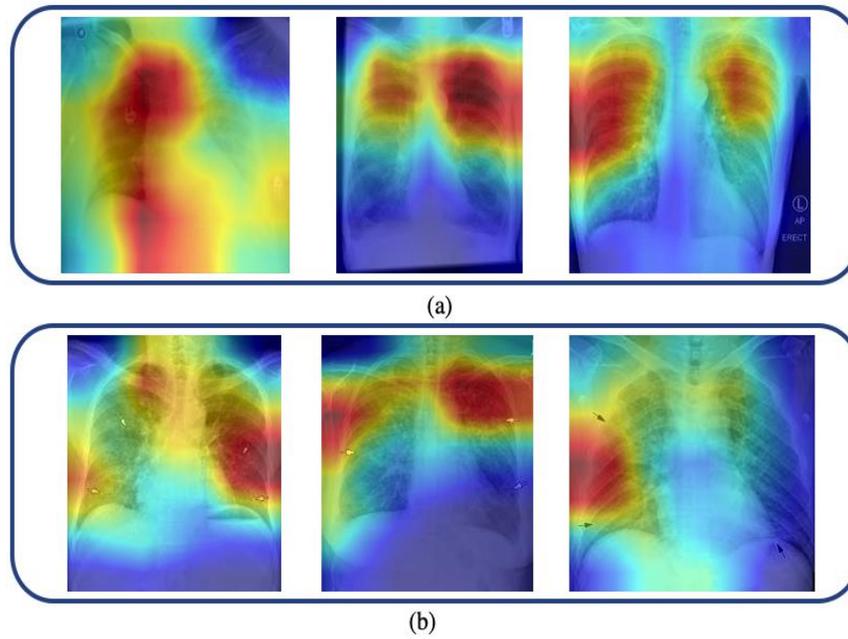
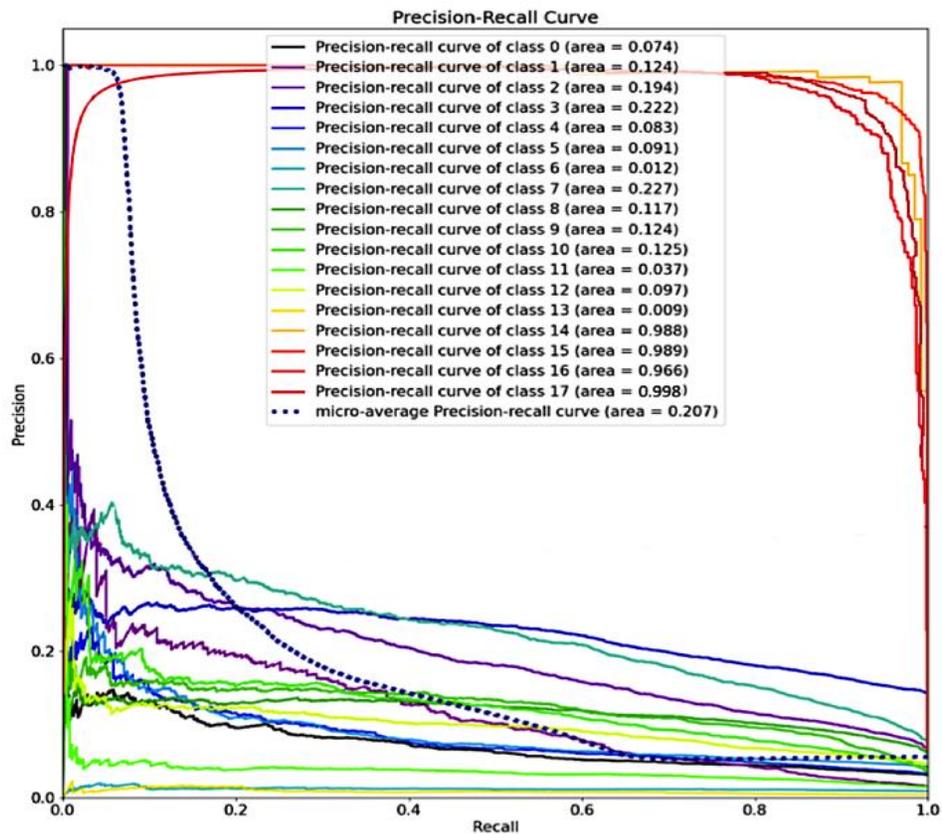
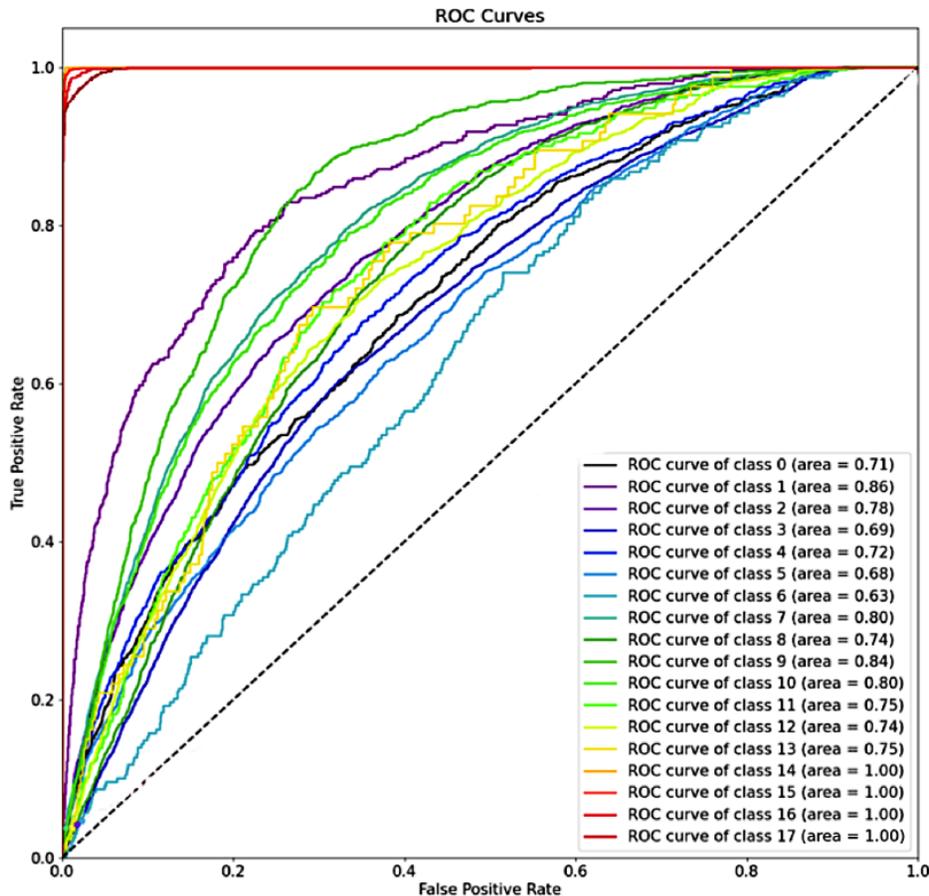


FIGURE 9. Precision-Recall of Each Class



On the basis of Figure 9, it is shown that the class curves of 14,15,16, and 17 achieve a balance between recall and precision of 0.998, 0.989, 0.966, and 0.999, respectively. Perfectly skilled models are represented as points at coordinates (1,1). Using the DenseNet121 model, the PR curve appears to produce the same results as the ROC curve.

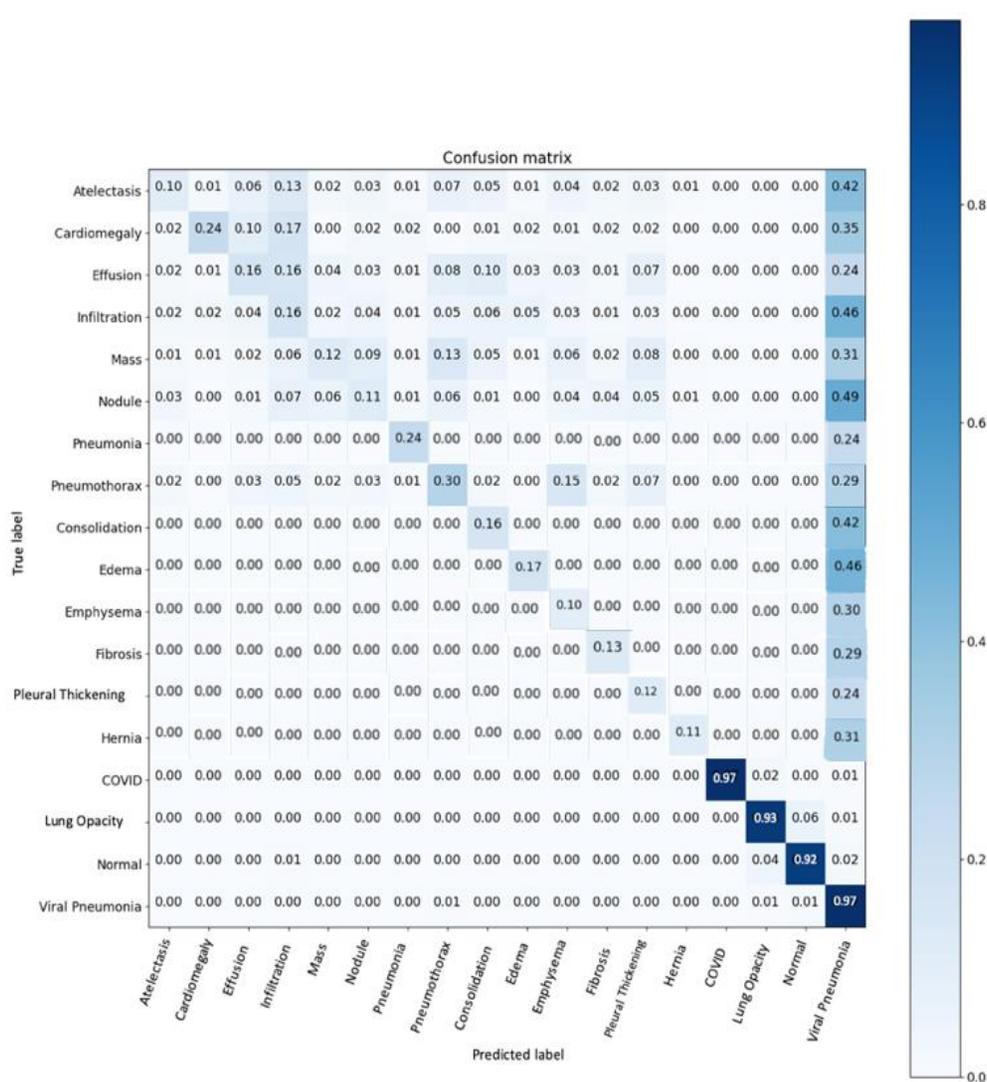
FIGURE 10. ROC of Each Class



The 18 curves in figure 10 demonstrate sequential colour degradation, from dark blue to dark red. By interpreting the image above, it can be concluded that the orange, orange, bright red and dark red curves have superior performances than the others, namely classes 14, 15, 16, and 17 (viral pneumonia, normal lung, lung opacity, and COVID areas). Using this approach, the DenseNet121 model can be used to determine whether or not the techniques used to solve cases of lung disease classification are appropriate. Additionally, if you want to compare the curve's performance value numerically, the area under the curve (AUC) can be used. Clearly, the AUC is

the same as that found in table 2 judging from Figure 10.

FIGURE 11. Confusion Matrix Result



Despite our proposed network architecture achieving high AUC values in several categories of the NIH Chest X-ray dataset and all categories of the COVID-19 Radiography dataset, its usefulness in a clinical setting depends greatly on the availability of training and evaluation data. Specifically, the NIH dataset has significant label noise and the manner in which the label is interpreted by doctors. It has been noted that in the ChestX-ray14 dataset, the class "pneumothorax" is sometimes used for cases that have already been treated (i.e. the image shows the drain used to treat the pneumothorax). According to the confusion matrix in Figure 11, many

classes in the NIH dataset are predicted to be "Viral Pneumonia" lung diseases. It can be stated that "Viral Pneumonia" corresponds to a variety of lung diseases.

## 6. DISCUSSION

Providing evidence for the effectiveness of CNN-based transfer learning with DenseNet121 and dual datasets to diagnose lung diseases is the main objective of this study. These architectures have the biggest benefit of alleviating the vanishing gradient problem, strengthening feature propagation, encouraging feature reuse, and reducing the number of parameters. Additionally, CNN architectures based on transfer learning have become increasingly popular because of their increased accuracy. Therefore, it is increasingly common for machine learning methods to be combined with different pre-trained models [49].

Deep learning generally has the disadvantage that generalizations are heavily reliant on training data. COVID-19 has infected millions of people around the world today. Therefore, it is not certain whether the proposed deep learning-based studies will be successful on CT images from a different patient. By using millions of images during the training process, this uncertainty can be overcome. A data augmentation technique can increase the number of CT images and the accuracy of classifying them. However, these results do not provide as much learning as real-life examples do. For a general and real success, there should be an increase in image data and training. Data limitations make it impossible to obtain a true picture of success. In addition to requiring a more powerful computer, adding more data increases the training time. Despite its high success, the DenseNet121 structure proposed in this study uses transfer learning as its infrastructure. Additionally, ANN-based segmentation performed in the application does not separate the lungs from each other. Although the extracted features more accurately represent the infection than the raw images, segmenting both lung images by region would provide a more accurate estimate of the infection.

A model that has overfitted may have learned the training dataset too well, despite the statistical noise or random fluctuations in the dataset. An overfitted model will have more difficulty

generalizing to new data the more specialized it becomes to the training data, increasing the generalization error. The performance of the model on the validation dataset can be used to measure the increase in generalization error. It has been shown that more and more specialized models require greater generalization errors as they become less able to generalize to new data as they become more specialized to training data. The performance of the model on the validation dataset can be used to measure the increase in generalization error. As a result, too much flexibility may result when the model has more capacity than what is needed for the problem. An overly long training period can also result in this. Training could be halted at the point at which the validation loss inflection point occurs, as the dynamics of overfitting are evident after that point [50], [51].

## 7. CONCLUSION

An automated analysis of chest X-ray data and a COVID-19 chest X-ray dataset was developed in this work to detect lung diseases. The X-ray images we used for this study were publicly available and consisted of 133,280. DenseNet121 trained on images  $224*224*3$  to ensure accurate classification. 82.16 % of lung diseases can be detected by this model. For more accurate identification of lung diseases, we intend to investigate chest X-rays and combine them with other models to perform experiments. In order to prevent the spread of lung disease to others, it is extremely important to detect lung disease as early as possible. COVID-19 chest X-rays were used with four classes and general chest X-rays with 14 classes in this study.

Having a DenseNet learning structure with the residual concept in it is the reason the proposed model has this learning structure. A chest X-ray image can be used to diagnose COVID-19 disease using an automatic model that allows for the detection of lung diseases without the need for handcrafted feature extraction techniques. Fast, stable, and easy-to-use software can therefore help the radiographer make better decisions. Consequently, the radiologists' workload is reduced, and misdiagnoses are prevented.

In future research, even though the dual dataset method is successful, a more advanced deep learning method will be proposed for lung disease detection. In the first study, more datasets will

be collected to increase efficacy. As is common knowledge, deep learning success is strongly influenced by the quantity of labeled data available. So this research will combine generative adversarial networks (GANs) with deep neural networks (DNNs). A study is also planned to develop a more robust CNN-based lung segmentation.

## **ACKNOWLEDGEMENTS**

We thank Faiz Ayyas Munawwar for providing the illustration in this paper.

## **SOURCE OF FUNDING**

This study is funded by Directorate General of Higher Education, Ministry of Education, Culture, Research, and Technology, Indonesia as a part of 2021 Applied Excellent Research in Higher Education Grant Number 163/E4.1/AK.04.PT/2021.

## **CONFLICT OF INTERESTS**

The author(s) declare that there is no conflict of interests.

## **REFERENCES**

- [1] D. R. Olson, M. Huynh, A. Fine, et al. Preliminary estimate of excess mortality during the COVID-19 outbreak - New York City, March 11-May 2, 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69:603–5.
- [2] Z. Rachik, S. Bidah, H. Boutayeb, O. Zakary, M. Rachik, Understanding the different objectives of information and their mutual impact: multi-information model, *Commun. Math. Biol. Neurosci.* 2021 (2021), Article ID 1.
- [3] H. T. Likassa, The impacts of covariates on spatial distribution of CORONA virus 2019 (COVID-19): WHO data through multivariate analysis of covariance (MANCOVA), *Eurasian J. Med. Oncol.* 4 (2) (2020), 141–148.
- [4] Y. Hu, J. Sun, Z. Dai, H. Deng, X. Li, Q. Huang, Y. Wu, L. Sun, Y. Xu, Prevalence and severity of corona virus disease 2019 (COVID-19): A systematic review and meta-analysis, *J. Clinic. Virol.* 127 (2020), 104371.
- [5] G. Radi, F. Diotallevi, A. Campanati, A. Offidani, Global coronavirus pandemic (2019-nCoV): implication for an Italian medium size dermatological clinic of a II level hospital, *J. Eur. Acad. Dermatol. Venereol.* 34(5) (2020),

e213–e214.

- [6] K.M. Agarwal, S. Mohapatra, P. Sharma, S. Sharma, D. Bhatia, A. Mishra, Study and overview of the novel corona virus disease (COVID-19), *Sensors Int.* 1 (2020), 100037.
- [7] N. Kong, C. Gao, M. S. Xu, Y. L. Xie, C. Y. Zhou, Spontaneous pneumomediastinum in an elderly COVID-19 patient: A case report, *World J. Clin. Cases*, 8 (2020), 3573–3577.
- [8] H. Chen, X. Wu, W. Wang, Q. Wang, When cancer encounters COVID-19 in China: what have we suffered, experienced and learned, *Japan. J. Clinic. Oncol.* 50 (2020), 712–717.
- [9] A. Syakriah, Terawan in hot water again, this time over radiology services, *The Jakarta Post*, (2020).
- [10] O. Primadi et al., *Profil Kesehatan Indonesia Tahun 2019*. (2020).
- [11] R. C. M. Yam, P. W. Tse, L. Li, P. Tu, Intelligent predictive decision support system for condition-based maintenance, *Int. J. Adv. Manuf. Technol.* 17 (2001), 383–391.
- [12] A. D. Revelos, *The evolution of radiology through product and process innovation*, PhD diss., Massachusetts Inst. Technol., pp. 1–195, 2015.
- [13] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, D. Shen, Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19, *IEEE Rev. Biomed. Eng.* 14 (2021), 4–15.
- [14] J. G. Lee, S. Jun, Y. W. Cho, H. Lee, G. B. Kim, J. B. Seo, N. Kim, Deep learning in medical imaging: general overview, *Korean J. Radiol.* 18 (2017), 570-584.
- [15] M. A. Bruno, E. A. Walker, H. H. Abujudeh, Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction, *Radiographics*, 35 (2015), 1668–1676.
- [16] R. C. of Radiologists, *Standards for the reporting and interpretation of imaging investigations*, 2006.
- [17] T. W. Cenggoro, B. Pardamean, PapSmear image recording system for artificial intelligence data collection, *IOP Conf. Ser.: Earth Environ. Sci.* 794 (2021), 012109.
- [18] D. Ardila, A.P. Kiraly, S. Bharadwaj, B. Choi, J.J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D.P. Naidich, S. Shetty, End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography, *Nat Med.* 25 (2019), 954–961.
- [19] B. J. Teng, M. F. Byrne, Introduction for the artificial intelligence and gastrointestinal cancer column, *J. Med.*

- Artif. Intell. 2 (2019), 14.
- [20] F. Valentino, T.W. Cenggoro, B. Pardamean, A design of deep learning experimentation for fruit freshness detection, *IOP Conf. Ser.: Earth Environ. Sci.* 794 (2021), 012110.
- [21] J.S. Suri, A. Puvvula, M. Biswas, et al. COVID-19 pathways for brain and heart injury in comorbidity patients: A role of medical imaging and artificial intelligence-based COVID severity classification: A review, *Computers Biol. Med.* 124 (2020), 103960.
- [22] X. Mei, H.-C. Lee, K. Diao, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19, *Nat. Med.* 26 (2020), 1224–1228.
- [23] N. Dominic, Daniel, T. W. Cenggoro, A. Budiarto, and B. Pardamean, “Transfer learning using inception-resnet-v2 model to the augmented neuroimages data for autism spectrum disorder classification, *Commun. Math. Biol. Neurosci.* 2021 (2021), 39.
- [24] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, *Proc. 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2261–2269, 2017.
- [25] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D. Jackel, Backpropagation applied to handwritten zip code recognition, *Neural Comput.* 1 (1989), 541–551.
- [26] R. E. Caraka, N. T. Nugroho, S. K. Tai, R. C. Chen, T. Toharudin, B. Pardamean, Feature importance of the aortic anatomy on endovascular aneurysm repair (EVAR) using Boruta and Bayesian MCMC, *Commun. Math. Biol. Neurosci.* 2020 (2020), Article ID 22.
- [27] B. Mahesworo, T. W. Cenggoro, A. Budiarto, F. R. Lumbanraja, B. Pardamean, Phosphorylation site prediction using gradient tree boosting, *Commun. Math. Biol. Neurosci.* 2020 (2020), Article ID 48.
- [28] L. Zhang, R. Rong, Q. Li, et al. A deep learning-based model for screening and staging pneumoconiosis. *Sci. Rep.* 2201 (2021), 11.
- [29] G. Chassagnon, M. Vakalopoulou, A. Régent, et al. Deep learning-based approach for automated assessment of interstitial lung disease in systemic sclerosis on CT images, *Radiology: Artificial Intelligence.* 2 (2020), e190006.
- [30] E. J. Hwang, S. Park, K. N. Jin, et al. Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs, *JAMA Netw. Open.* 2 (2019), e191095.
- [31] S. Bharati, P. Podder, M.R.H. Mondal, Hybrid deep learning for detecting lung diseases from X-ray images,

- Inform. Med. Unlocked. 20 (2020), 100391.
- [32] J. Shuja, E. Alanazi, W. Alasmay, A. Alashaikh, COVID-19 open source data sets: a comprehensive survey, *Appl. Intell.* 51 (2021), 1296–1325.
- [33] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, M. Ghassemi, COVID-19 image data collection: prospective predictions are the future, *J. Mach. Learn. Biomed. Imaging*, 2020 (2020), 2.
- [34] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.
- [35] A. Krizhevsky, One weird trick for parallelizing convolutional neural networks, *ArXiv:1404.5997 [Cs]*. (2014).
- [36] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2818–2826, 2016.
- [37] B. Pardamean, T. W. Cenggoro, R. Rahutomo, A. Budiarto, E. K. Karuppiah, Transfer learning from chest X-Ray pre-trained convolutional neural network for learning mammogram data, *Procedia Comput. Sci.* 135 (2018), 400–407.
- [38] K. Sriporn, C. F. Tsai, C. E. Tsai, P. Wang, Analyzing lung disease using highly effective deep learning techniques, *Healthcare*. 8 (2020), 107.
- [39] R. Takahashi, T. Matsubara, K. Uehara, Data augmentation using random image cropping and patching for deep CNNs, *IEEE Trans. Circuits Syst. Video Technol.* 30 (2020), 2917–2931.
- [40] T. D. Pham, Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning?, *Health Inf. Sci. Syst.* 9 (2021), 2.
- [41] Y. X. Tang, Y. B. Tang, Y. Peng, K. Yan, M. Bagheri, B.A. Redd, C.J. Brandon, Z. Lu, M. Han, J. Xiao, R.M. Summers, Automated abnormality classification of chest radiographs using deep convolutional neural networks, *Npj Digit. Med.* 3 (2020), 70.
- [42] F. Rahmad, Y. Suryanto, K. Ramli, Performance comparison of anti-spam technology using confusion matrix classification, *IOP Conf. Ser.: Mater. Sci. Eng.* 879 (2020), 012076.
- [43] O. S. Sushkova, A. A. Morozov, A. V. Gabova, A. V. Karabanov, S. N. Illarioshkin, A statistical method for exploratory data analysis based on 2D and 3D area under curve diagrams: Parkinson’s disease investigation, *sensors*. 21 (2021), 4700.

- [44] J. Mallick, S. Talukdar, M. Alsubih, Mohd. Ahmed, A.R.M.T. Islam, Shahfahad, N.V. Thanh, Proposing receiver operating characteristic-based sensitivity analysis with introducing swarm optimized ensemble learning algorithms for groundwater potentiality modelling in Asir region, Saudi Arabia, *Geocarto International*. (2021) 1–28.
- [45] N. V. Sarlis, E. S. Skordas, S.-R.G. Christopoulos, P. A. Varotsos, Natural time analysis: the area under the receiver operating characteristic curve of the order parameter fluctuations minima preceding major earthquakes, *Entropy*. 22 (2020), 583.
- [46] J. Cook, V. Ramadas, When to consult precision-recall curves, *Stata J*. 20 (2020), 131–148.
- [47] Encyclopedia, Emphysema, Encyclopedia, 2018. [Online]. Available: <https://www.encyclopedia.com/medicine/diseases-and-conditions/pathology/emphysema>. [Accessed: 15-Sep-2020].
- [48] M. Carol DerSarkissian, What is viral pneumonia?, Web MD, 2020. [Online]. Available: <https://www.webmd.com/lung/viral-pneumonia>. [Accessed: 15-Sep-2020].
- [49] D. Wang, J. Mo, G. Zhou, L. Xu, Y. Liu, An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images, *PLoS ONE*. 15 (2020), e0242535.
- [50] N.E.M. Khalifa, M.H.N. Taha, A.E. Hassanien, S. Elghamrawy, Detection of coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset, *ArXiv:2004.01184 [Cs, Eess]*. (2020).
- [51] N.-A.-A. Alam, M. Ahsan, Md.A. Based, J. Haider, M. Kowalski, COVID-19 detection from chest X-ray images using feature fusion and deep learning, *Sensors*. 21 (2021), 1480.